Extended version of the article published in the
International Journal on Artificial Intelligence Tools (2008).

PREDICTIONS AND DIAGNOSTICS IN EXPERIMENTAL DATA USING SUPPORT VECTOR REGRESSION

NIKITA A. SAKHANENKO

Computer Science Department, University of New Mexico
Albuquerque, NM 87131 USA
sanik@cs.unm.edu

GEORGE F. LUGER

Computer Science Department, University of New Mexico
Albuquerque, NM 87131 USA
luger@cs.unm.edu

HANNA E. MAKARUK

E548 Physics Division, Los Alamos National Laboratory
Los Alamos, NM 87545 USA
hanna_m@lanl.gov

DAVID B. HOLTKAMP

D410 Physics Division, Los Alamos National Laboratory
Los Alamos, NM 87545 USA
dholtkamp@lanl.gov

In this paper we present a novel support vector machine (SVM) based framework for prognosis and diagnosis. We apply the framework to sparse physics data sets, although the method can easily be extended to other domains. Experiments in applied fields, such as experimental physics, are often complicated and expensive. As a result, experimentalists are unable to conduct as many experiments as they would like, leading to very unbalanced data sets that can be dense in one dimension and very sparse in others. Our method predicts the data values along the sparse dimension providing more information to researchers. Often experiments deviate from expectations due to small misalignments in initial parameters. It can be challenging to distinguish these outlier experiments from those where a real underlying process caused the deviation. Our method detects these outlier experiments. We describe our success at prediction and outlier detection and discuss implications for future applications.

Keywords: Support vector machines; regression; sparse data; prediction; diagnosis.

1. Introduction

In this paper we describe a powerful new approach to prediction, diagnosis, and global data-integrative analysis using SVM regression. We demonstrate the power of our approach by analyzing a high-energy physics environment. Our tool, based on the SVM technology, is initially applied to velocity data obtained during shock

physics experiments on tin. However, as our results indicate, the same method can easily be adapted to many other applied physics problems, including experiments on different types of metal, under different physical conditions, and even with significantly different set-up, leading the way for analysis of other related applications.

Experimental physics, along with many other fields in applied research, uses experiments, physical tests, and observations to gain insight into various phenomena as well as to validate hypotheses and models. Shock physics is a field that explores the response of materials to the extremes of pressure, deformation, and temperature which are present when shock waves interact with those materials.[1] High explosives (HE) are often used to generate these strong shock waves. Many different diagnostic approaches have been used to probe these phenomena.[2]

Because of the energetic nature of the shock wave drive, often a large amount of experimental equipment is destroyed during the test. Similar to many other applied sciences, the cost and complexity of repeating a significant number of experiments, or conducting a systematic study of some physical property, are simply too costly to conduct to the degree of completeness and detail that researchers desire. As a result, a data analyst is often left with data sets whose dimensions are highly diverse: a data set might be very dense in one dimension and sparse in another. Additionally, high energy released during the experiment contributes noise, quickly increasing as the equipment degrades. Finally, since each test requires multiple parameters of the physical system to be fine-tuned, physicists often encounter various data misalignment issues when attempting to interpret the results.

Our approach utilizes a novel variant of support vector machine learning that interpolates the shock physics data along the sparse data dimension. The method supplies physicists with new indirect information that is implicit when traditional data analysis is used. Moreover, our approach allows for prediction of the physical measurements under new experimental conditions without repeating a necessary set of costly experiments. Predictability of the data from the experiments by itself provides more insight about the underlying physical process. Furthermore, we also focus on the problem of identifying "outlier" experiments, i.e., those experiments that for some reason went wrong. Our method can diagnose which experimental data do not fit with data sets from other "good" experiments. Another important application of the method is for comparison and integration of the predicted information with other kinds of data, including those from simulated models.

In section 2 we further describe the details and tasks of our SVM-based method. In section 3 we illustrate the method by applying it to a physics example whose results are evaluated in section 4. Finally, we provide related work and conclude.

2. SVM-Based Prognosis and Diagnosis

2.1. *Support vector machines*

The Support Vector Machine (SVM) uses *supervised learning* to estimate a functional input/output relationship from a set of training data. Formally, given the

training data set of k points $\{\langle x_i, y_i\rangle | x_i \in X, y_i \in Y, i = 1, \ldots, k\}$, that is independently and randomly generated by some unknown function f for each data point, the Support Vector Machine method finds an approximation of the function, assuming f is of the form

$$f(x) = w \cdot \phi(x) + b, \tag{1}$$

where ϕ is a nonlinear mapping $\phi : X \to H$, $b \in Y$, $w \in H$. Here $X \subseteq R^n$ is the input space, $Y \subseteq R$ is the output space, and H is a high-dimensional feature space. The coefficients w and b are found by minimizing the *regularized risk*[3]

$$R = C \sum_{i=1}^{k} L(f(x_i), y_i) + \lambda \parallel w \parallel^2 . \tag{2}$$

This formula shows that regularized risk R consists of an empirical risk, defined via a loss function, complemented with a regularization term. In this paper we measure the empirical risk using an ε-intensive loss function[4] L defined as

$$L(f(x), y) = \begin{cases} |f(x) - y| - \varepsilon, \text{ if } |f(x) - y| \geq \varepsilon \\ 0, \qquad\qquad\quad \text{otherwise} \end{cases} .$$

Minimizing the regularization term, $\lambda \parallel w \parallel^2$, enforces the resulting function to be as flat as possible, hence controlling how general the function is (which is very crucial in extremely noisy domains). Constant C in (2) is called a *regularization constant* or a *capacity factor*, and ε is the size of the ε-tube (also called an *error-insensitive zone* or an *ε-margin*). Note that ε determines the accuracy of the regression, namely the amount by which a point from a training set is allowed to diverge from the regression. Note also that the support vector machine is a method involving kernels. Recall that the kernel of an arbitrary function $g : X \to Y$ is an equivalence relation on X:

$$ker(g) = \{(x_1, x_2) | x_1, x_2 \in X, g(x_1) = g(x_2)\} \subseteq X \times X.$$

We can think of a kernel as a nonlinear similarity measure.

Originally, the SVM technique was applied to classification problems, in which the algorithm finds the maximum-margin hyperplane in the transformed feature space H that separates the data into two classes. The result of an SVM used for regression estimation (Support Vector Regression, SVR) is a model that depends only on a subset of training data, because the loss function used during the modeling omits the training data points inside the ε-tube (points that are close to the model prediction).

The SVM approach has several attractive features pointed out by Shawe-Taylor and Cristianini.[5] One of these features is the good generalization performance which an SVM achieves by using a unique principle of structural risk minimization.[6] In addition, SVM training is equivalent to solving a linearly constrained quadratic programming problem that has a unique and globally optimal solution, hence there is no need to worry about local minima. A solution found by SVM depends only on

a subset of training data points, called *support vectors*, making the representation of the solution sparse.

Finally, since the SVM method involves kernels, it allows us to deal with arbitrary large feature spaces without having to compute explicitly the mapping ϕ from the data space to the feature space, hence avoiding the need to compute the product $w \cdot \phi(x)$ of Eq.(1). In other words, a linear algorithm that uses only dot products can be transformed by replacing dot products with a kernel function. The resulting algorithm becomes non-linear, although it is still linear in the range of the mapping ϕ. We do not need to compute ϕ explicitly, because of the application of kernels. This algorithm transformation from the linear to non-linear form is known as a *kernel trick*.[7]

2.2. *The tasks*

In applied fields, such as experimental physics, a data set consists of information obtained from a number of various tests. Often researchers cannot conduct as many experiments as necessary to complete a study, due to complexity and cost of these tests. The first problem we consider in this paper is to predict the measurement values for missing experiments.

Furthermore, we also focus on experiments whose data recordings were not all successful. The configuration of the experiments in applied fields is controlled by multiple parameters whose precise calibration is very crucial for a successful test. Even a small deviation in any of these parameters as well as in hidden environmental variables can set the experiment off, making its results less informative. Often only domain experts with a lot of experience can immediately distinguish between such "outlier" experiments and tests with "real" physical phenomena. Our method can provide a diagnostic function to experimentalists detecting the "outlier" experiments and possibly identifying which parameters of the test system went wrong.

The task of increasing the informational output of experimental data is important, due to the limited number of experiments, their difficult implementation, and high cost. Researchers, who attempt to explain all the phenomena of these experiments, can gain more insight by combining the experimental data with our SVM-based predictions. Moreover, the predicted model can further support and even improve the understanding of other types of data obtained during the experiment. Another important application of SVM-based data estimations is for comparison with various kinds of numerical experiments (in physics these called *hydrocode models*) generated by large programs that simulate various hard or impossible to perform experiments. Consequently, the model can be incorporated into a data manifold of the experiment data.

2.3. *Equivalent problem*

Consider each data point from a given data set as a tuple $\langle t_1, \ldots, t_n \rangle$ for some $n > 2$, where each t_i represents a recorded data value in the ith direction. One can

see that these data points lie on a surface in the n-dimensional space. Hence the problem identified in section 2.2 can be transformed into the task of reconstructing the surface from the given data.

In other words, the problem is to find a regression of one coordinate on the rest of the coordinates of a sample based on data. Formally, consider n random variables, $T_1, \ldots, T_n$. The problem is to estimate coefficients $\theta \in \Theta$ such that the error

$$e = T_1 - \rho(T_2, \ldots, T_n; \theta) \tag{3}$$

is small, where ρ is a regression function, that is, $\rho : \mathbb{R}^{n-1} \times \Theta \to \mathbb{R}$, and $\Theta \subseteq \mathbb{R}$ is a set of coefficients of the model. Note that variables $T_2, \ldots, T_n$ are the $n-1$ factors of a regression, and T_1 is an observation.

3. The Example: an Application in Experimental Physics

We next illustrate our SVM-based method by applying it to the analysis of surface velocity data taken from HE shocked tin samples using a laser velocity interferometer called a VISAR.[8,9,10] These experiments have been described elsewhere in detail.[11] For the purposes of this paper, it is sufficient to note that the VISAR data presented here describe the response of the free surface of the metal coupon to the shock loading and HE generated shock wave. Physicists analyze the time dependence of the velocity magnitude to obtain information on the yield strength of the material, and the thickness of the leading damage layer that may separate from the bulk material during the experiment.

3.1. *The experiment*

The data were obtained from experiments in which metal samples are damaged/melted during a high explosive detonation with single point ignition. A schematic view of the experiment setup is shown in figure 1. A cylindrically shaped metal coupon is positioned on top of a 12.7 mm thick high explosive (HE) disk. Both the metal and HE coupons are 50.8 mm in diameter. A point detonator is glued to the center of the HE disc in order to perform single point ignition symmetrically. Note that all of the components of the experiment setup have a common axis of symmetry for consistent data analysis. During an experiment a VISAR probe, located on the axis above the metal sample, transmits a laser beam, and the velocity of the top surface of the metal is inferred from the light reflected from the coupon. The time series of the velocity measured throughout an experiment constitutes the VISAR velocimetry. Note that there are other types of data obtained during each experiment,[11] this paper is devoted to the analysis of VISAR velocimetry data.

There are two parameters that vary between different experiments: the metal type of the sample and the thickness of the coupon. By changing the thickness of the metal coupon and the type of metal in the initial setup of an experiment, researchers attempt to see the changes in physical processes across the set of experiments. Only the experiments on tin samples are described in this paper.

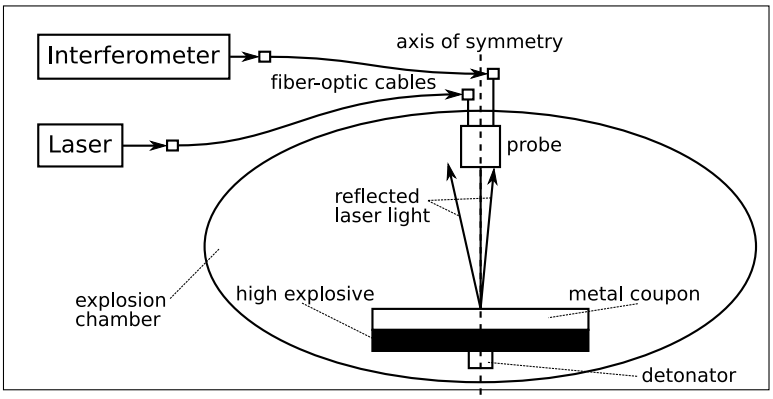

Fig. 1. Schematic representation of the experiment setup.

3.2. *VISAR data*

A Velocity Interferometer System for Any Reflector (VISAR) is a system designed to measure the Doppler shift of a laser beam reflected from the moving surface under consideration so as to capture changes of the velocity of the surface. The VISAR system is able to detect very small velocity changes (a few meters per second when the velocity could be more than a thousand meters per second). Moreover, it is able to measure even the velocity changes of a diffusely reflecting surface.

A VISAR system consists of lasers, optical elements, detectors, and other components as shown in figure 1. The light is delivered from the laser via optical fiber to the probe and is focussed in such a way that some of the light reflects from the moving surface back to the probe. The reflected laser light is transmitted to the interferometer. Note that since the reflected light is Doppler shifted, the interferometer extracts the velocity of the moving surface from the wavelength change of the light.

This method, widely used in the experiments similar to the one described in section 3.1, is reasonably reliable. For instance, the measurements obtained using a VISAR system are in agreement with the results obtained by Makaruk et al.[12] Since the method of information extraction proposed by Makaruk et al. is independent of VISAR, it additionally validates VISAR results.

Since the available data are the VISAR measurements that capture some characteristics of the unknown function, and each data point is represented by several features, the data are suitable for the application of supervised learning methods, such as SVR. A velocity of each data point is a target value for SVR, whereas the thickness and time are feature values. In other words, random variable T_1 from (3) represents velocity, and T_2 and T_3 stand for thickness and time ($n = 3$).

In figure 2 we describe the VISAR data set. It is important to note that the data are significantly stretched along the time dimension. This happens because

the whole dataset is comprised of time series corresponding to a set of measured experiments. During each experiment, the VISAR readings were recorded every 2ns for as long as 6000 time steps. However, for some experiments the VISAR system finished recording useful information earlier than for other experiments. The data were cut by the shortest sequence (1656 time steps), since it has been identified experimentally that SVM performs better on the aligned data. On the other hand, if we consider VISAR measurements across the thickness dimension, the data cover the thicknesses starting from 6.35 mm up to 12.7 mm with 1.5875 mm increases. In total, 5 time sequences of 1656 points comprise the data used by the SVM method.

Figure 2 presents the complete data set projected on the $Time \times Velocity$ plane. The original data set is represented by dotted lines and is smoothed using a sliding triangular window, which is depicted by solid lines. The amount of the time steps, where each step is equal to 2ns, is shown on the abscissa.

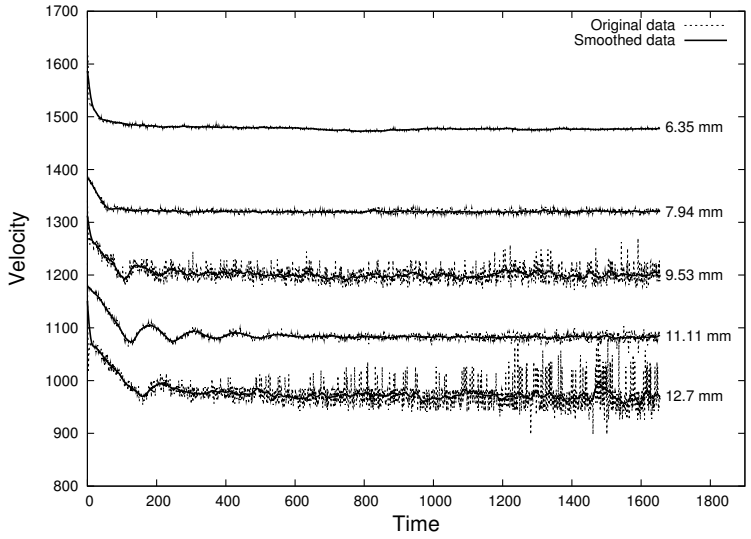

Fig. 2. The projection of the VISAR data set and its smoothed version.

In order to identify the best application of the SVM technology to the VISAR data, we used k-fold cross-validation. The data are divided into k parts, out of which $k-1$ parts are used for training the learning machine, and the last part is used for its validation. The process is repeated k times using each part of the partitioning precisely once for validation.

4. Evaluation of Results

There are several factors affecting the quality of the resulting regression analysis. The error of VISAR data as well as the errors occurring during the data prepro-

cessing affect the accuracy of the reconstructed surface. It is generally agreed[8,9,10] that a VISAR system measures the velocity values with an absolute accuracy of 3-5%. This is an approximate error calculated from differences between repeated experiments. Although the number of repeated experiments was too small to allow a more robust statistical analysis, this level of uncertainty is in the range of values generally agreed on by VISAR experimenters.[8,9,10] Measurement error, together with noise, transfers into the regression result. In addition, since the ignition time (the start of the experiment) was different with different experiments, data have to be time-aligned so as to make each time series start from the moment of the detonation. This introduces another potential error into the regression.

The accuracy of the reconstructed surface is also affected by the specific features of VISAR data. The length of each of the time series produced by the VISAR system during different experiments always differs. We have observed that the SVM performs better on the data combined from the time series of the same length than from those of different length. Hence, the length of the data was aligned. In addition, each data point of three elements (velocity, time, and thickness) has order 10^3, 10^{-6}, and 1. This is why it is important to scale the data to improve the performance of the SVM.

Unfortunately, the application of SVR directly to the set of smoothed and aligned data yields overfitted results. This overfit results from the data step in the time direction being much smaller than the step in the other directions; hence for any chosen data range there are more data points along the time axis than along the thickness axis. The overfitting problem is solved by scaling the data in such a way that the distance between two neighbor points along any axis is equal to 1.

Using nonlinear kernels achieves better performance when the dynamics of an experiment are non-linear. It is known that Gaussian Radial Basis Function (RBF) kernels perform well under general smoothness assumption,[13] hence a Gaussian RBF

$$k(x, y) = e^{-\gamma \|x-y\|^2}$$

was chosen as the kernel for the reconstruction. Additionally, it has been experimentally determined that SVM techniques with simpler kernels, such as polynomial, take longer to train and return non-satisfactory results.

The performance of the SVR with RBF kernel was directly affected by three parameters, the radius γ of RBF, the regularization constant C, and the size ε of the ε-tube which determines the accuracy of the regression (see section 2.1). k-fold cross-validation was performed in order to determine the optimal parameter values under which SVR produces the best approximation of the surface. An l^2 error is computed for each parameter instantiation after finishing the cross-validation. Figure 3 demonstrates how the error changes depending on the values of the SVR parameters.

It can be seen in figure 3 that the error increases as the radius γ goes up. The error also increases when ε becomes bigger. One can also see that the change of

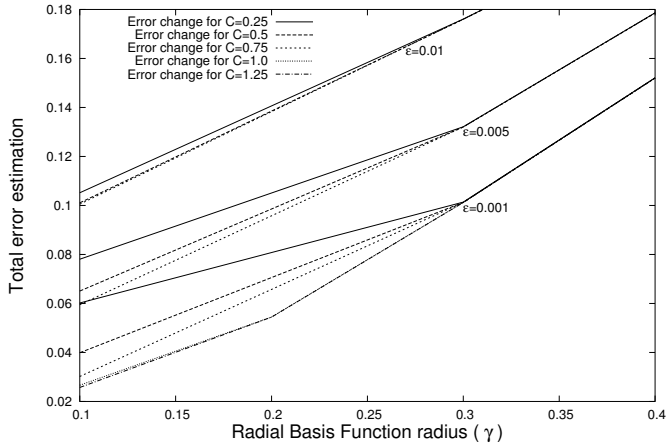

Fig. 3. Error changes depending on different model parameters.

C affects the error the most when γ is the smallest, and the influence of C on the error decreases as γ goes up, becoming insignificant when γ exceeds 0.3. At the same time, given a small γ, parameter C affects the error more as ε decreases. The error analysis suggests that when the tuple $\langle \gamma, C, \varepsilon \rangle$ is around $\langle 0.1, 0.75 - 1.0, 0.001 \rangle$, the total error is minimized. This error analysis produces a range of suboptimal values for the parameters. Expert knowledge is used in order to identify the final model that returns the most accurate velocity surface, shown in figure 4.

When this surface is found, it is possible to predict a velocity value for any $\langle time, thickness \rangle$ pair. Once the surface is accurate and stable enough, VISAR data that deviate significantly from the surface can be identified consequently detecting the "outlier" experiments. The surface provides significantly more information about velocity changes across the thickness dimension than do the VISAR readings alone. It can also provide velocity time series for an experiment in which only imagery data were measured, successfully improving the quality of the analysis for this experiment, and, consequently increasing the understanding of the whole physical system.

In this paper we used an implementation of SVM regression techniques called *SVM-light*. For more information about implementation details see (Ref. 14).

5. Related Work

The SVM technology is used for both classification and regression tasks. Most of the various applications of SVM are for classification, including handwriting recognition,[15] face detection in images[16] and lip tracking in video,[17] speech recognition[18] and speech emotion detection,[19] and various pattern recognition tasks in bioinformatics.[20,21] Some of the applications of SVM in physics include the work

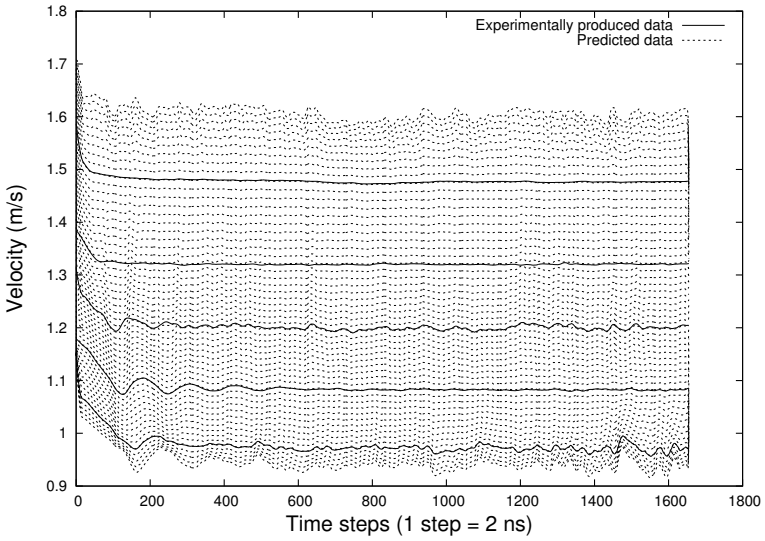

Fig. 4. SVM prediction results: dotted lines represent the prediction of the time series for the thicknesses between those that are produced experimentally (the solid lines).

by Vannerem et al.[22] testing SVM in the physics environment by using support vector classifiers in the analysis of simulated high energy physics data, and by Cai et al.[23] presenting another example of the use of SVM techniques in the analysis of physics data when the SVM is used to classify sonar signals.

In the case of regression, SVMs have been applied to financial forecasting,[24] superresolution problems in image processing,[25] benchmark time series prediction tests,[26] stream flow data estimation,[27] and regularization of model inversion.[28] Our work is very different from financial and other time series forecasting[24,26] using SVMs for regression, because we essentially predict the data values *between* time series as opposed to predicting the values at the next time steps. Research of Dibike et al.[27] who successfully applied SVMs for regression to the problem of stream flow data estimation based on records of rainfall and other climatic data, is related to our research on prediction of velocity data from time and thickness parameters. On the other hand, our approach provides an outlier-experiment detection tool as well as produces rich information suitable for integration into a data manifold of the physical experiment. To our knowledge this is the first attempt to use support vector regression for detection and data integration.

6. Conclusions and Future Work

In this paper we described the tool for data prediction along the sparse dimension of the data set. Our method is based on support vector regression reconstructing a data surface in the data space. We applied our method to VISAR velocity data ob-

tained from high-energy physics experiments and successfully estimated the velocity surface in $Time \times Thickness \times Velocity$ data space. The optimization parameters of the method are obtained using cross-validation and grid search and then further validated by the domain expert. In case of velocity data prediction, our method provides considerably more information about the velocity behavior as a function of time and thickness than experimentally produced VISAR measurements alone. This, in turn, significantly improves the scientific value of VISAR data in other areas of analysis of physics experiments, such as in proton radiography imagery analysis[11,12] and in computational simulations.

Since it is based on SVM, our method does not require a vast amount of data for producing good data estimations. This is very helpful when used in applied fields where available data are limited due to the high cost and complexity of experiments. In addition, we show that our method can be used for outlier experiment detection, i.e., it can be used to distinguish between experiments with intrinsic underlying governing process and experiments that significantly deviated due to disarrangement of system's parameters and other factors. Application in experimental physics revealed this as an important advantage of our tool, since often a lot of domain knowledge is needed to identify the outlier experiments.

There are several future directions of our work. One of these is to investigate the possibility of using a custom kernel instead of the standard Gaussian. Intuitively, an elliptical kernel that accounts for the high density of data in one direction and sparsity in all other directions may improve the results of the method applied to a unbalanced data set such as the velocity data considered in this paper. Investigation of different techniques for the search of SVM free parameter values, such as online learning algorithms for SVM parameter fitting, is another direction for further research.

Finally, note that our method used for prediction produces a point estimate. However, most of the time we wish to capture uncertainty in the prediction, hence estimating the conditional distribution of the target values given feature values is more attractive. There are a number of different extensions to the SVM technique and hybrids of SVM with Bayesian methods, such as *relevance* vector machines and Bayesian SVM, that use probabilistic approaches.[29,30] Exploring these methods could give significantly more information about the underlying data.

Acknowledgments

The authors thank Joysree Aubrey of Los Alamos National Laboratory, LANL, for numerous long and thought-provoking discussions. Special thanks to Brendt Wohlberg of LANL for providing ideas about SVM applications. This work was supported by the Department of Energy under the ADAPT program.

References

1. Ya. B. Zel'dovich, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, (Dover Publications, Mineola, NY, 2002).
2. W. M. Isbell, *Shock Waves: Measuring the Dynamic Response of Materials*, (Imperial College Press, London, 2005).
3. B. Schölkopf and A. Smola, *Learning with Kernels. Support Vector Machines, Regularization, Optimization, and Beyond*, (MIT Press, 2001).
4. V. Vapnik, *The Nature of Statistical Learning Theory*, (Springer, 1999).
5. J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*, (Cambridge University Press, 2004).
6. V. Vapnik, *Estimation of dependences based on empirical data*, (Springer Verlag, New York, 1982).
7. M. Aizerman, E. Braverman and L. Rozonoer, Theoretical foundations of the potential function method in pattern recognition learning, in *Automation and Remote Control*, Vol. 25 (1964), pp. 821–837.
8. L. M. Barker and R. E. Hollenback, Laser Interferometer for Measuring High Velocities of any Reflecting Surface, in *Journal of Applied Physics*, Vol. 43 (1972), pp. 4669–4675.
9. L. M. Barker and K. W. Schuler, Correction to the Velocity-per-Fringe Relationship for the VISAR Interferometer, in *J. of Applied Physics*, Vol. 45 (1974), pp. 3692–3693.
10. W. F. Hemsing, Velocity Sensing Interferometer (VISAR) Modification, in *Rev. Sci. Instrum*, Vol. 50 (1979), pp. 73–78.
11. D. B. Holtkamp et al., A Survey of High Explosive-Induced Damage and Spall in Selected Materials Using Proton Radiography, in *Shock Compression of Condensed Matter*, AIP Conference Proceedings, Vol. 706 (2003), pp. 477–482.
12. H. E. Makaruk, M. A. Sakhanenko and D. B. Holtkamp, Analysis of Proton Radiography Images of Shock Melted/Damaged Tin, in *Tech. Report* LA-UR-06-107, (Los Alamos National Laboratory, 2006).
13. A. Smola, B. Schölkopf and K.-R. Müller, The connection between regularization operators and support vector kernels, in *Neural Networks*, Vol. 11 (1998), pp. 637-649.
14. T. Joachims, Making large-Scale SVM Learning Practical, in *Advances in Kernel Methods – Support Vector Learning*, ed. B. Schölkopf et al. (MIT Press, 1999).
15. C. J. C. Burges and B. Schölkopf, Improving the accuracy and speed of support vector learning machine, in *Advances in neural information processing systems*, Vol. 9 (MIT Press, Cambridge, MA, 1997), pp. 375–381.
16. M. Türkan, B. Dülek, I. Onaran and A. E. Cetin, Human face detection in video using edge projections, in *Proc. of SPIE-6246*, (2006).
17. R. Y. D. Xu, A Portable and Low-Cost E-Learning Video Capture System, in *Advanced Concepts for Intel. Vision Systems*, Vol. 4179 (Springer LNCS, 2006), pp. 1088–1098.
18. W. M. Campbell, J. P. Campbell, D. A. Reynolds, E. Singer and P. A. Torres-Carrasquillo, Support vector machines or speaker and language recognition, in *Computer and Speech Language*, Vol. 20 (2006), pp. 210–229.
19. Y.-H. Cho, K.-S. Park and R. J. Pak, Speech Emotion Pattern Recognition Agent in Mobile Communication Environment Using Fuzzy-SVM, in *Fuzzy Information and Engineering*, Vol. 40 (Springer Advances in Soft Computing, 2007), pp. 419–430.
20. T. S. Furey, N. Duffy, N. Cristianini, D. Bednarski, M. Schummer and D. Haussler, Support Vector Machine Classification and Validation of Cancer Tissue Samples Using Microarray Expression Data, in *Bioinformatics*, Vol. 16(10) (2000), pp. 906–914.
21. P. P. Luedi, A. J. Hartemink, R. L. Jirtle, Genome-wide prediction of imprinted murine genes, in *Genome Res.*, Vol. 15 (2005), pp. 875–884.

22. P. Vannerem, K.-R. Müller, B. Schölkopf, A. Smola and S. Söldner-Rembold, Classifying LEP data with Support Vector algorithms, in *AIHENP* (To be published, 1999).
23. C.-Z. Cai, W.-L. Wang, and Y.-Z. Chen, Support Vector Machine classification of physical and biological datasets, in *International Journal of Modern Physics C*, Vol. 14(5) (2003), pp. 575–585.
24. T. B. Trafalis and H. Ince, Support vector machine for regression and applications to financial forecasting, in *Proc of IJCNN*, (2000), pp. 348–353.
25. K. S. Ni and T. Q. Nguyen, mage superresolution using support vector regression, in *IEEE Trans. on Image Processing*, Vol. 16(6) (2007), pp. 1596–1610.
26. L. Cao, Support vector machines experts for time series forecasting, in *Neurocomputing*, Vol. 51 (2003), pp. 321–339.
27. Y. B. Dibike, S. Velickov, D. Solomatine and M. B. Abbott, Model induction with Support Vector Machines: introduction and applications, in *Journal of Computing in Civil Engineering*, Vol. 15(3) (2001), pp. 208–216.
28. S. S. Durbha, R. L. King and N. H. Younan, Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer, in *Remote Sensing of Environment*, Vol 107 (2007), pp. 348–361.
29. M. E. Tipping, The Relevance Vector Machine, in *Advances in Neural Information Processing Systems*, (Morgan Kaufmann, San Mateo, CA, 2000).
30. C. M. Bishop and M. E. Tipping, Variational Relevance Vector Machine, in *Proc. of UAI*, ed. C. Boutilier and M. Goldszmidt (Morgan Kaufmann, 2000), pp. 46–53.